Do Free Trade Agreements Actually Increase Members' International Trade?

Scott L. Baier
The John E. Walker Department of Economics
Clemson University
Clemson, SC 29634
and the Federal Reserve Bank of Atlanta, Atlanta, GA 30309
sbaier@clemson.edu

Jeffrey H. Bergstrand*
Department of Finance
Mendoza College of Business
and Kellogg Institute for International Studies
University of Notre Dame
Notre Dame, IN 46556
bergstrand.1@nd.edu

Abstract

For over 40 years, the gravity equation has been a workhorse for cross-country empirical analyses of international trade flows and – in particular – the effects of free trade agreements (FTAs) on trade flows. However, the gravity equation is subject to the same econometric critique as earlier cross-industry studies of U.S. tariff and nontariff barriers and U.S. multilateral imports: trade policy is *not* an exogenous variable. We address econometrically the endogeneity of FTAs using instrumental-variables (IV), control-function (CF), and panel-data techniques; IV and CF approaches do not adjust for endogeneity well, but a panel approach does. Accounting econometrically for the FTA variable's endogeneity yields striking empirical results: the effect of FTAs on trade flows is *quintupled*.

October 2004

*Corresponding author. The authors are grateful to Jonathan Eaton, Rob Feenstra, Scott Kastner, Mika Saito, Curtis Simon, Jeff Wooldridge, Eduardo Zambrano, two anonymous referees and participants at presentations at the American Economic Association annual meetings, Midwest International Economics meetings, and ISNIE meetings for excellent comments on earlier drafts. We thank Doru Cojuc, Katie Martin, Kushlani Rajapakse, Duncan Stewart and Li Zhang for excellent research assistance. The authors are grateful to the National Science Foundation for financial support.

Do Free Trade Agreements Actually Increase Members' International Trade?

> The issue of exogeneity may also be an important problem when dummy variables are used (in a gravity equation) to estimate the effects of free trade areas (Lawrence, 1998, p. 59).

One might expect – having witnessed a virtual explosion in the number of regional free trade agreements (FTAs) among nations over the past decade and a half – that the answer to the question posed in this paper's title is unequivocal: *yes*! Surprisingly, international trade economists can actually claim little firm empirical support for reliable quantitative estimates of the average effect of an FTA on bilateral trade (all else constant).

Over 40 years, the "gravity equation" has emerged as the empirical workhorse in international trade to study the *ex post* effects of FTAs and customs unions on bilateral merchandise trade flows.[1] The gravity equation is typically used to explain cross-sectional variation in country pairs' trade flows in terms of the countries' incomes, bilateral distance, and dummy variables for common languages, for common land borders, and for the presence or absence of an FTA. Nobel laureate Jan Tinbergen (1962) was the first to publish an econometric study using the gravity equation for international trade flows, which included evaluating the effect of FTA dummy variables on trade. His results suggested economically insignificant "average treatment effects" of FTAs on trade flows. Tinbergen found that membership in the British Commonwealth (Benelux FTA) was associated with only *5 (4) percent* higher trade flows. Since then, results have been mixed, at best. Aitken (1973), Abrams (1980), and Brada and Mendez (1983) found the European Community (EC) to have an economically and statistically significant effect on trade flows among members, whereas Bergstrand (1985) and Frankel, Stein and Wei (1995) found insignificant effects. Frankel (1997) found positive significant effects from Mersosur, insignificant effects from the Andean Pact, and significant *negative* effects from membership in the EC:

> If the data from four years – 1970, 1980, 1990, 1992 – are pooled together, the estimated coefficient on the European Community is a smaller 0.15, implying a 16 percent effect" (p. 83).

Frankel (1997) concludes that several readers "have found surprising our result that intra-European trade can be mostly explained by various natural factors, with little role for the EC until the 1980s . . . ."(p. 88).

---

[1]Bayoumi and Eichengreen (1997, p. 142) note that the gravity equation has "long been the workhorse for empirical studies of the pattern of trade." In 1992, the European Union commissioned a major study to analyze the potential enlargement of the union into central and eastern Europe. That study, now Baldwin (1994), used the gravity equation as the primary empirical tool. The gravity equation has also been used to study the effects of partial preferential arrangements, cf., Sapir (1981); however, we will not address those here. This study (purposefully) does not address *ex ante* analyses of the effects of FTAs on trade flows using computable general equilibrium models.

Other studies in international trade have had similar seemingly implausible results.[2]

The fragility of estimated FTA treatment effects is addressed directly in Ghosh and Yamarik (2004). These authors use extreme-bounds analysis to test the robustness of FTA dummy coefficient estimates. They find empirical evidence using cross-section data that the estimated average treatment effects of most FTAs are "fragile," supporting our claims. Thus, there still are no reliable *ex post* estimates of the FTA average treatment effect (ATE). This paper is aimed at addressing this puzzle.

All these studies, however, typically assume an *exogenous* RHS dummy variable to represent the FTA treatment effect. In reality, FTA dummies are not exogenous random variables; rather, countries likely select endogenously into FTAs, perhaps for reasons unobservable to the econometrician and possibly correlated with the level of trade.[3] This paper applies developments in the econometric analysis of treatment effects – some well-known and others more recent – to estimate the effects of FTAs on bilateral trade flows using a panel of cross-section time-series data at five-year intervals from 1960 to 2000 for 96 countries. The literature on treatment effects, developed in the context of numerous labor economics studies (cf., Wooldridge, 2002), provides rich tools that have not previously been used for analyzing the effects of bilateral trade policies on international trade flows.

This is not the first paper in empirical international trade to call attention to the potential endogeneity bias in estimating the effect of trade policies on trade volumes. For instance, Trefler (1993) addressed systematically the simultaneous determination of U.S. multilateral imports and U.S. multilateral nontariff barriers in a cross-industry analysis. Trefler found using instrumental variables that, after accounting for the endogeneity of trade policies, the effect of these policies on U.S. imports increased *tenfold*. Lee and Swagel (1997) also showed using instrumental variables that previous estimates of the impact of trade liberalization on imports had been considerably underestimated.

Clearly, the literature on bilateral trade flows and bilateral FTAs using the gravity equation is subject to the same critique that Trefler raised: the presence or absence of an FTA is *not exogenous*. The issue is important because – if FTAs are endogenous – previous cross-section and panel empirical estimates of the effects of FTAs on trade flows may be biased and inconsistent, and the effects of FTAs

---

[2] Frankel (1997) and Oguledo and MacPhee (1994) provide summaries of FTA coefficient estimates across studies. Frankel (pp. 86-90) draws considerable attention to the surprising insignificant effects (especially prior to the 1980s) of the EC and EFTA in his and others studies, such as Bergstrand (1985, 1989) and Boisso and Ferrantino (1997). No systematic explanation is provided.

[3] We note that, for about a decade, several researchers have acknowledged potential endogeneity bias, but only that created by GDPs as RHS variables. Several authors have instrumented for GDPs, but (with the exception of the two studies noted shortly) none have instrumented for FTAs.

on trade may be seriously over- or under-estimated, as the extreme-bounds evidence in Ghosh and Yamarik (2004) suggest. To date, only two papers have attempted to address the bias in cross-section gravity models caused by endogenous FTAs, Baier and Bergstrand (2002) and Magee (2003). However, both papers – using instrumental variables – provide at best mixed evidence of isolating the effect of FTAs on trade flows.

The empirical results in our paper suggest three important conclusions. First, several plausible reasons exist to suggest that the quantitative (long-run) effects of FTAs on trade flows using the standard cross-section gravity equation are *biased*; we argue that unobservable heterogeneity most likely biases estimates downward. Second, we find that, owing to this bias, traditional estimates of the effect of FTAs on bilateral trade flows have tended to be underestimated by as much as *75-85 percent*. Third, we demonstrate in this paper – using recent theoretical models and econometric advances – that the best method to estimate the effect of FTAs on bilateral trade flows employs a theoretically-motivated gravity equation using *differenced panel data*, rather than instrumental variables applied to cross-section data.

Section 1 presents the (more traditional) atheoretical cross-section gravity equation and a (modern) theoretically-motivated cross-section gravity equation. Section 2 provides motivation for suspecting endogeneity of FTA dummy variables and suggests the likely direction of bias. Section 3 addresses instrumental-variable and control-function econometric methods for eliminating endogeneity bias and discusses empirical results using cross-section data. Section 4 addresses panel techniques and discusses the results from applying fixed effects and first-differencing to panel data. Section 5 concludes.

1. The Gravity Equation in International Trade

The gravity equation in international trade most commonly estimated using cross-country data is:

$$PX_{ij} = \beta_0 (GDP_i)^{\beta_1} (GDP_j)^{\beta_2} (DIST_{ij})^{\beta_3} e^{\beta_4 (LANG_{ij})} e^{\beta_5 (ADJ_{ij})} e^{\beta_6 (FTA_{ij})} \epsilon_{ij} \tag{1}$$

where PX_{ij} is the value of the merchandise trade flow from exporter i to importer j, GDP_i (GDP_j) is the level of nominal gross domestic product in country i (j), $DIST_{ij}$ is the distance between the economic centers of countries i and j, $LANG_{ij}$ is a binary variable assuming the value 1 if i and j share a common language and 0 otherwise, ADJ_{ij} is a binary variable assuming the value 1 if i and j share a common land border and 0 otherwise, FTA_{ij} is a binary variable assuming the value 1 if i and j have a free trade agreement and 0 otherwise, e is the natural logarithm base, and ϵ_{ij} is assumed to be a log-normally distributed error term.[4]

The earliest applications of the gravity equation to international trade flows were not grounded in formal theoretical foundations, cf., Tinbergen (1962), Linnemann (1966), Aitken (1973) and Sapir (1981); these earlier studies appealed either to informal economic foundations or to a physical science analogy. Since 1979, however, formal theoretical economic foundations for a gravity equation similar to equation (1) have surfaced, cf., Anderson (1979), Bergstrand (1985), Deardorff (1998), Baier and Bergstrand (2001), Eaton and Kortum (2002), and Anderson and van Wincoop (2003). A notable feature common to *all* these models is a role for *prices*; in all six papers, price levels or some form of multilateral price indexes surface theoretically.[5] Most recently, Anderson and van Wincoop (2003) illustrate succinctly the omitted variables bias introduced by ignoring prices in the cross-section gravity equation. Their framework suggests theoretically the gravity model should be estimated as:

$$\ln[X_{ij} /(GDP_i \, GDP_j)] = \beta_0 + \beta_3 \ln DIST_{ij} + \beta_4 (ADJ_{ij}) + \beta_5 (LANG_{ij}) + \beta_6 (FTA_{ij}) - \ln P_i^{1-\sigma} - \ln P_j^{1-\sigma} + \epsilon_{ij} \tag{2}$$

subject to N equilibrium conditions:

$$P_1^{1-\sigma} = \sum_{i=1}^{N} P_i^{\sigma-1} (GDP_i / GDP^W) \, e^{\beta 3 \ln DISTi1 + \beta 4 (LANGi1) + \beta 5 (ADJi1) + \beta 6 (FTAi1)} \tag{3.1}$$

$$\vdots$$

$$P_N^{1-\sigma} = \sum_{i=1}^{N} P_i^{\sigma-1} (GDP_N / GDP^W) \, e^{\beta 3 \ln DISTiN + \beta 4 (LANGiN) + \beta 5 (ADJiN) + \beta 6 (FTAiN)} \tag{3.N}$$

to estimate β_0, β_3, β_4, β_5 and β_6. GDP^W denotes world GDP, constant across countries. $P_i^{1-\sigma}$ and $P_j^{1-\sigma}$ are

---

[4]The standard gravity equation sometimes includes the exporter and importer populations or per capita GDPs. Theoretical foundations for this alternative specification still require refinement, which is beyond the scope of this particular paper but an issue addressed in another paper of the authors. Nevertheless, the results of this paper hold up when populations are added; these complementary results are omitted for brevity.

[5]A seminal contribution to the theoretical foundations of the gravity equation, but in the absence of trade costs (and consequently price terms), is Helpman and Krugman (1985).

denoted "multilateral (price) resistance terms."[6] Anderson and van Wincoop then estimate this system using a custom nonlinear least squares program, treating all $P_i^{1-\sigma}$ variables (i=1,...,N countries) as endogenous. However, Anderson and van Wincoop (2003) and Feenstra (2004, Ch. 5) both suggest that an alternative – and computationally easier – method for accounting for multilateral price terms $P_i^{1-\sigma}$ and $P_j^{1-\sigma}$ in cross section – that will also generate unbiased coefficient estimates of β_0, β_3, β_4, β_5 and β_6 – is estimation of equation (2) using country-specific fixed effects.[7]

As preliminary empirical support for our claim that FTA coefficients are biased, we estimate a typical cross-section gravity equation for multiple years: 1960, 1970, 1980, 1990, and 2000. Table 1 provides cross-section coefficient estimates for these years of a typical (log-linear version of) gravity equation (1), estimated using the (non-zero) nominal trade flows among 96 countries identified in Appendix 1. The data sources are standard and are provided in section 3.[8] The five sets of estimates in Table 1 support our claims and findings in the literature that the FTA dummy's coefficient estimates are highly unstable from year to year, are often small positive values, and in some years are even *negative*.[9]

Table 2 provides estimates for the same years of gravity equation (2), where – following Anderson and van Wincoop (2003, Table 6) and Feenstra (2004, Ch. 5) – we use country fixed effects to account for the "multilateral price" terms (rather than a custom nonlinear least squares program). The FTA dummy coefficient estimates remain unstable across years and, for several years, the FTA coefficient's estimate is negative.[10] While the fixed effects help to account for the endogeneity bias

---

[6]Assuming distance, language, adjacency and *FTA* are symmetric for *ij* and *ji*, the equilibrium multilateral resistance terms can be solved for according to equations (3.1)-(3.N).

[7]Eaton and Kortum (2002), Rose and van Wincoop (2001), and Redding and Venables (2000) similarly used country fixed effects to account for multilateral price terms.

[8]The issue of non-zero trade flows has been treated in other papers, such as Eichengreen and Irwin (1995) and Felbermayr and Kohler (2004). However, this issue is beyond the scope of the present paper. For instance, in 2000, we have 9120 potential trade flows among 96 countries (96x95=9120), less 1573 trade flows recorded as zero and 245 observations recorded as "missing."

[9]The basic findings in Table 1 (and Table 2) are generally unaffected when we considered fixed sample sizes for all five years. First, we restricted our sample to all trading partners who had positive trade flows in 1960; for any of these pairs that had zeros in 1970, ..., 2000, we substituted ones for zeros. Second, we restricted our sample to all trading partners who had positive trade flows in 2000; for any of these pairs that had zeros in 1960, ..., 1990, we substituted ones for zeros. The coefficient estimates were materially the same across samples.

[10]The reader may also notice the upward trend (in absolute values) in non-*FTA* coefficient estimates with time in Tables 1 and 2. This also is an interesting issue but is beyond the scope of this paper; the upward drift (in absolute value) of the distance coefficient estimate has been studied by other researchers, cf., Felbermayr and Kohler (2004). Our focus is instead on the instability of the *FTA* coefficient estimates. Indeed, these coefficient estimates are much more unstable. For instance, using Table 2's specification, the coefficients of variation of the coefficient estimates for the non-*FTA* variables ranged from 0.29 to 0.31. By contrast the coefficient of variation of the *FTA* coefficient estimates was 2.05, six times that of the others.

created by prices and the influence of FTAs among other countries on the trade from i to j, they do not correct for the bias introduced if countries select into FTAs. We must turn to other methods to identify unbiased FTA treatment effects, which is the purpose of this paper.

2. Endogeneity Bias

A standard problem in cross-section empirical work is the potential endogeneity of right-hand-side (RHS) variables. If any of the RHS variables in equation (1) or (2) are correlated with the error term, ϵ_{ij}, that variable is considered econometrically "endogenous" and OLS may yield biased and inconsistent coefficient estimates. Potential sources of endogeneity bias of RHS variables generally fall under three categories: omitted variables, simultaneity, and measurement error (see Wooldridge, 2002, pp. 50-51). While we believe all three factors may contribute potentially to endogeneity bias caused by *FTA*, we will argue that the most important source is omitted variables (and selection) bias. We discuss each source in turn.

2.A. Omitted Variables (and Selection) Bias

The gravity equation has surfaced over the past four decades as the dominant empirical framework for analyzing trade flows primarily because of its strong explanatory power. Explanatory power (R^2) generally ranges from 60 to 80 percent. However, as Table 1 suggests, there remains considerable unobserved heterogeneity among country pairs (leaving aside the issue of multilateral price terms, for now).

In determining the potential correlation between the gravity equation's error term, ϵ_{ij}, with FTA_{ij}, one first needs to consider what determines the likelihood of an FTA between a pair of countries. Although trade economists have examined empirically for many years the determinants of tariff rates and non-tariff barrier levels across industries and across countries, virtually no empirical work has examined the determinants of FTA_{ij}. A notable exception is Baier and Bergstrand (2004). These authors present a theoretical and empirical model of economic determinants of FTAs. Their paper finds strong empirical evidence that pairs of countries that have FTAs *tend* to share economic characteristics that their theory suggests should enhance economic welfare of the pairs' representative consumers. For instance, two countries tend to have an FTA the larger and more similar their GDPs, the closer they are to each other but the more remote the pair is from the rest-of-the-world (ROW), and the wider (narrower) the difference in their relative factor endowments with respect to each other (the ROW). But these include the same factors that tend to explain large trade flows. Thus, in terms of *observable* economic characteristics, countries with FTAs have "chosen well," in the sense that most country pairs with FTAs tend to have the economic characteristics associated with considerable trade and with (in theory) welfare-enhancing net trade creation. Yet, the estimated probit functions in Baier and Bergstrand (2004) have pseudo-R^2 values of only 70 percent, still leaving considerable unobserved heterogeneity.

The important question for this paper is: How is the unobserved heterogeneity in trade flow determinants associated with the decision of whether or not to form an FTA? For instance, error term ϵ_{ij} in equations (1) or (2) may be representing unobservable (to the econometrician) policy-related barriers – tending to reduce trade between two countries – that are not accounted for by standard gravity equation RHS variables. As an example, suppose two countries have extensive unmeasurable domestic regulations (e.g., internal shipping regulations) that inhibit trade (causing ϵ_{ij} to be negative). The likelihood of the two countries' governments selecting into an FTA may be high if there is a large expected welfare gain from potential bilateral trade creation if the FTA deepens liberalization beyond tariff barriers into domestic regulations (and other non-tariff barriers). Thus, FTA_{ij} and the intensity of domestic regulations may be positively correlated, but the gravity equation error term ϵ_{ij} and the intensity of domestic regulations may be negatively correlated. This reason suggests that FTA_{ij} and ϵ_{ij} are negatively correlated, and the *FTA* coefficient will tend to be *underestimated*.

In support of this argument, numerous authors have noted that one of the major benefits of regionalism is the potential for "deeper integration." Lawrence (1996, p. xvii) distinguishes between "international policies" that deal with border barriers, such as tariffs, and "domestic policies" that are concerned with everything "behind the nation's borders, such as competition and antitrust rules, corporate governance, product standards, worker safety, regulation and supervision of financial institutions, environmental protection, tax codes ..." and other national issues. The GATT and WTO have been remarkably effective in the post-WWII era reducing border barriers. However, these institutions have been much less effective in liberalizing the domestic policies just named. As Lawrence states it, "Once tariffs are removed, complex problems remain because of differing regulatory policies among nations" (p. 7). He argues that in many cases, FTA "agreements are also meant to achieve deeper integration of international competition and investment" (p. 7). Gilpin (2000) echos this argument: "Yet, the inability to agree on international rules or to increase international cooperation in this area has contributed to the development of both managed trade *and regional arrangements*" (p. 108; italics added). Preeg (1998) concludes:

> [Free] trade agreements over time, however, have tended to include a broader and broader scope of

> other trade-related policies. This trend is a reflection, in part, of the fact that as border restrictions [tariffs] are reduced or eliminated, other policies become relatively more important in influencing trade flows and thus need to be assimilated in the trade relationship (p. 50).

We believe this omitted variable bias is the major source of endogeneity facing estimation of FTA effects in gravity equations. With cross-section data, standard econometric techniques to address omitted variables (and selection) bias include estimation using instrumental variables and Heckman control functions. Alternatively, with panel data, fixed effects and first differencing can be employed to treat endogeneity bias. We discuss these approaches in sections 3 and 4, respectively.

2.B. Simultaneity Bias

Consider the potential endogeneity bias created by simultaneity. GDP – a function of exports and imports – is potentially endogenous to bilateral trade flows. Yet, three convincing reasons exist for largely ignoring potential endogeneity of incomes. First, GDP is a function of *net* exports, not gross exports. Typically, net exports tend to be less than 5 percent of a country's GDP (in absolute terms). Second, the gravity equation relates *bilateral* trade flows to countries' incomes. Trade between any pair of countries tends to be a very small share of any country's exports, much less its GDP. Third, Frankel (1997) and others have previously accounted for the potential endogeneity of national incomes econometrically using instrumental variables (IV) including labor forces and stocks of human and physical capital. Frankel (1997) reported that coefficient estimates in gravity equations change insignificantly using these IV techniques and concluded "Evidently, the endogeneity of income makes little difference" (p. 135). Nevertheless, we can still account easily for the potential endogeneity of GDPs using a regression specification such as equation (2) with GDPs on the LHS.

Of the remaining RHS variables in equation (1), only *FTA* seems potentially endogenous. It is plausible to treat *DIST* and *ADJ* as exogenous. Moreover, the presence or absence of a common language (*LANG*) may reasonably be treated in cross-section as exogenous also.

Yet, as discussed earlier, there exists a large empirical literature in international trade on the effects of multilateral tariff and nontariff barriers on multilateral trade volumes, and the simultaneous effects of these trade volumes on multilateral barriers using cross-industry and cross-country data for particular years, c.f. Trefler (1993) and Swagel and Lee (1997). Simultaneity may be an issue for *FTA* in cross-section gravity equations, motivated as in these two studies. For example, holding constant typical gravity equation RHS variables (*GDPs, DIST, ADJ, LANG*), two countries (say, the United States and China) that possibly trade more than their "natural" level, as predicted by the typical gravity equation, may create political pressures to avoid trade liberalization or possibly raise trade barriers. This would cause a negative simultaneity bias in the *FTA* coefficient estimate. On the other hand, the governments of two countries that trade more than their gravity-equation-suggested "natural" level might be induced to form an FTA because there might potentially be less "trade diversion" due to their extensive trading relationship, suggesting a positive simultaneity bias.

To address this issue, the natural inclination is to estimate a system of simultaneous equations treating bilateral trade and FTAs as endogenous variables using 2SLS, as in Magee (2003) following in the spirit of Trefler (1993) and Lee and Swagel (1997). However, as proved in Heckman (1978) and Maddala (1983, p. 118), the system of equations estimated in Magee (2003) is not logically consistent.[11] Moreover, as shown in Angrist and Krueger (2001, p. 80), the simultaneous-equations approach using instrumental variables taken in Magee (2003) will yield inconsistent coefficient estimates in the second-stage gravity equation *unless* the (first-stage) probit function "happens to be exactly right." The details are provided in Appendix 2.

2.C. Measurement Error Bias

Finally, measurement error in an explanatory variable, such as an FTA dummy, is generally associated with negative bias (in absolute terms) in the variable's coefficient. For instance, with the classical "errors-in-variables" assumption, the 0-1 FTA dummy variable (*FTA*) would be correlated positively with the measurement error (ς) if the true trade-policy variable (say, the tariff rate, t) was assumed uncorrelated with ς ($\varsigma = FTA - t$). In equation (1)'s context, the correlation between *FTA* and the error term ($\epsilon - \beta_6 \varsigma$) would be negative, leading to the classical "attenuation bias" of *FTA*'s coefficient estimate toward zero.[12] We believe that this may be part of the reason – but neither the entire, nor even the most important, reason – *FTA* coefficient estimates have been underestimated.

Of course, the best method for eliminating this bias is construction of a continuous variable that would more accurately measure the degree of trade liberalization from various FTAs. If FTAs only eliminated bilateral tariff rates, one would ideally measure this liberalization with a change in the *ad valorem* tariff rate (for which data is poor). However, FTAs liberalize trade well beyond the elimination

---

[11] Magee's approach is analogous to that in Trefler (1993) and Lee and Swagel (1997), with the notable exception that the latter two studies both involved instead two *continuous* endogenous variables in a system of *linear* equations. The econometric issues differ.

[12] Even without the classical errors-in-variable assumption, the correlation between *FTA* and ϵ is likely negative. Suppose the true trade policy variable is the bilateral tariff rate, t, where $t > 0$. If an FTA exists, $FTA = 1$, $t = 0$, and ς consequently equals 1. If no FTA exists, $FTA = 0$, $t > 0$, and consequently $\varsigma < 0$. Thus, *FTA* and ς are positively correlated.

of tariffs. Calculation of such measures is beyond the scope of this particular study, but is a useful direction for future research. Our goal rather is to determine reliable estimates of the *treatment effect* of an FTA, similar to the 0-1 variable representing program participation in labor econometrics. Thus, we constrain our study to estimate more accurately the *ex post* effect of an FTA dummy on trade flows, as has been employed in the gravity equation literature for over four decades.[13]

3. Treatment Effects in a Cross-Section Model of Trade Flows and Free Trade Agreements

This section has three parts. In the first part, we address conventional cross-section instrumental-variables (IV) and control-function approaches to addressing the FTA endogeneity bias associated with omitted variables and selection. In the second part, we discuss data issues. In the third part, we provide estimates of FTA treatment effects using IV and control-function methods.

3.A. Cross-Section Econometric Techniques

Historically, the gravity equation in international trade has been estimated using cross-section data either for a single year or for multiple years; only very recently has the gravity equation been estimated using panel data techniques. In this section, we address the methods available for adjusting for the endogeneity bias of FTAs in gravity equations using cross-section data owing to omitted variables; in section 4, we address methods using panel data.

Methodological issues regarding the cross-section estimation of the partial effects of an endogenous binary variable (such as *FTA*) on a continuous endogenous variable (such as trade flows) fall under the "treatment effect" literature in econometrics. The average treatment effect refers to the notion that the trade flow between two countries will differ depending upon whether the countries share an FTA or not. The fundamental econometric dilemma is that one can observe only one situation or the other. An excellent summary of developments in the treatment-effect literature is in Heckman (2001) and Wooldridge (2002).

The average treatment effect (ATE) of an FTA between a country pair is defined as:

$$ATE(\boldsymbol{q}) \equiv E(x_1 - x_0 \mid \boldsymbol{q}, FTA) \tag{4}$$

where x_1 (x_0) denotes the logarithm of the sum of the trade flows between countries i and j with (without)

13 Another interesting direction suggested by one referee is to also examine the treatment effects by individual agreements. However, to limit the scope of this paper, this useful direction is left for future research.

the FTA, E denotes the expectation operator, and observation subscripts (ij) are omitted for simplicity. We assume that the level of trade between two countries depends upon the array of exogenous "covariates" *other than* the treatment – the standard set of gravity equation variables in equation (1) in levels or log levels as appropriate, excluding *FTA* – which we denote by $\boldsymbol{q}$ (*GDPs, DIST, ADJ, LANG*). Initially, we assume that observations of x are independently and identically distributed across country pairs; this assumption ensures that the treatment of one pair does not affect another pair's trade flow.[14] [15]

Consistent estimation of the average treatment effect depends upon assumptions made about relationships among the variables. Of course, in reality we do not observe both x_1 and x_0. Since one can only observe a trade flow in the presence *or* absence of an FTA, we define the observed outcome (x) for a country pair as:

$$x \equiv (FTA)\, x_1 + (1 - FTA)\, x_0 \tag{5}$$

where $FTA = 1$ if an FTA exists between the pair, and 0 otherwise. If we assume trade flows x_0 and x_1 have the standard linear form as in a gravity equation, then:

$$x_0 = \mu_0 + \beta'\boldsymbol{q} + \epsilon_0 \tag{6}$$
$$x_1 = \mu_1 + \beta'\boldsymbol{q} + \epsilon_1 \tag{7}$$

Substituting equations (6) and (7) in equation (5) yields:

$$x = \mu_0 + \beta'\boldsymbol{q} + \alpha FTA + \epsilon_0 + FTA(\epsilon_1 - \epsilon_0) \tag{8}$$

where $\alpha = \mu_1 - \mu_0$ corresponds to the average treatment effect.

Consistent estimation of the parameters in equation (8) depends upon correlations between the variables and the error terms. Specifically, consistent estimation of α, β, and μ_0 depends on the correlation of: (i) *FTA* with error term ϵ_0, and (ii) *FTA* with differences in unobservables for partners with FTAs versus partners without FTAs ($\epsilon_1 - \epsilon_0$). The former is the correlation associated with omitted variables bias, and the latter with selection bias. When *FTA* is uncorrelated with both factors, the parameters are estimated consistently using OLS.

[14] This is a strong assumption because it suggests that an FTA between countries i and j does not have a trade-diverting effect on other bilateral relations, such as countries i and m or j and n. However, we maintain this assumption only initially; later in this paper we account explicitly for it using multilateral price terms.

[15] We can also estimate the average treatment effect on countries that have formed an FTA, $TTE(\boldsymbol{q}, FTA=1) \equiv E(spx_1 - spx_0 \mid \boldsymbol{q}, FTA=1)$. This is known as the "treatment effect on the treated" (TTE), but is beyond the scope of this paper.

Consider first the case where *FTA* is correlated with ϵ_0, but $\epsilon_1 - \epsilon_0 = 0$ (for all country pairs). Wooldridge (2002) suggests an efficient IV estimator. First, assume the probability of *FTA* (P*(FTA)*) can be estimated by a known parametric form (such as probit) such that $P(FTA = 1 \mid \boldsymbol{q}, \boldsymbol{z}) = \Phi(\pi_0 + \pi_1'\boldsymbol{q} + \pi_2'\boldsymbol{z})$ where $P(FTA = 1 \mid \boldsymbol{q}, \boldsymbol{z}) \neq P(FTA = 1 \mid \boldsymbol{q})$ and $\boldsymbol{z}$ is a set of exogenous variables (not in $\boldsymbol{q}$).[16] Second, assume the variance of ϵ_0 is a constant. This suggests a multi-step IV method with the following steps: (i) estimate a binary response model, such as probit, $P(FTA = 1 \mid \boldsymbol{q}, \boldsymbol{z}) = \Phi(\pi_0 + \pi_1'\boldsymbol{q} + \pi_2'\boldsymbol{z})$ by maximum likelihood to generate predicted probabilities Φ^P; and (ii) estimate equation (8) by IV, using instruments 1, Φ^P, $\boldsymbol{q}$ and $\boldsymbol{z}$. The IV estimator is consistent, asymptotically efficient, and the usual 2SLS standard errors and test statistics are asymptotically valid.[17]

If *FTA* is correlated with ϵ_0 and $\epsilon_1 - \epsilon_0$, Heckman (1997) suggests that IV estimation only identifies the parameters consistently when economic agents' decisions to select into a program are unrelated to (or ignore) unobservable factors influencing the outcome. However, as discussed earlier, many trade-policy analysts have noted that policies tending to inhibit trade, such as nontariff barriers and domestic regulations, may be one of the main reasons governments have selected into FTAs. The intent of forming an FTA is that this agreement will lead to bilateral reductions in domestic barriers that multilateral agreements have been unable to attain. In the context of this study, IV estimation will not yield consistent estimates in the presence of selection bias. In this case, we apply Heckman's procedure to control for selection. For brevity, the estimation technique and associated empirical results are provided in Appendix 3.

3.B. Data

As established in the previous section, the key to estimating a consistent ATE of FTAs on trade is, first, finding a probit function that predicts FTAs and, second, finding variables for $\boldsymbol{z}$ (instruments) that are uncorrelated with the gravity equation error term. Before examining the possible instruments, we describe briefly the data used for the gravity equation. Nominal bilateral trade flows are from the International Monetary Fund's *Direction of Trade Statistics* for the years 1960, 1965, . . ., 2000 for 96 (potential) trading partners (zero trade flows are excluded); these data are deflated by GDP deflators to generate real trade flows for the panel analysis later. Nominal GDPs are from the World Bank's *World*

---

[16] If a probit function, then $\Phi(\ )$ is the standard normal cumulative distribution function.

[17] Wooldridge (2002) refers to this as a two-step estimator. However, for clarity, we refer to three steps. The first stage is the estimation of the predicted probabilities, Φ^P. The second stage is a linear regression of *FTA* on a constant, Φ^P, $\boldsymbol{q}$ and $\boldsymbol{z}$. The third stage is estimation of the gravity equation substituting the predicted values from the second-stage regression for *FTA*.

Development Indicators (2003); these will be scaled later by GDP deflators for the panel analysis. Bilateral distances were compiled using the *CIA Factbook* for longitudes and latitudes of economic centers to calculate the great circle distances. The FTA dummy variable was calculated using appendices in Lawrence (1996) and Frankel (1997), various websites, and FTAs notified to the GATT/WTO under GATT Articles XXIV or the Enabling Clause for developing economies; we included only full (no partial) FTAs and customs unions. Table 3 lists the trade agreements used and sources.[18] Capital-labor ratios were calculated using data from Baier, Dwyer, and Tamura (2003), which draw upon Mitchell (1992, 1993, 1995).

We will introduce several "political" variables as possible instruments that might be correlated with two governments' decision to form an FTA, but might *not* be correlated with their trade flows *per se*. Three themes underlie the variables' selection. First, there is now some evidence from the political science literature that FTAs are more likely to be formed when the countries' governments are more "democratic." Mansfield, Milner, and Rosendorff (2002) develop a theoretical and empirical model that shows that, as a government becomes more democratic, its leaders derive higher benefits from cooperating with other governments on liberalizing commercial transactions and entering FTAs. Forming FTAs conveys positive information to voters about their leaders fostering international cooperation, which tends to increase their reelection probabilities. Second, the formation of an FTA requires negotiations among governments. An FTA is more likely to be formed the lower the "costs" of bilateral trade-policy negotiations. These costs might be lower the less the regulatory intrusion of the governments and the more credible are the governments in committing to policies. Third, a large political science literature argues that FTAs generate a "security" externality, cf., Gowa and Mansfield (1993). Nations' governments can internalize this externality by forming an FTA with an ally.

We capture these three potential political influences with several variables. We use a standard indicator of average "democracy" of a pair of countries described in Jaggers and Gurr (1999). We supplement this democracy variable with some aggregate "governance indicators" from the World Bank that are potentially correlated with citizens' democratic rights, cf., Kaufmann, Kraay, and Mastruzzi (2003), averaged over the years 1996-2000. We use two indices of the process by which governments are selected, monitored and replaced ("Voice and Accountability" and "Political Stability and Absence of Violence") to represent citizens' democratic rights. The Voice-and-Accountability indicator measures the extent to which citizens of a country are able to participate in the selection of its government. The

[18] The data set is available at the authors' websites (http://www.nd.edu/~jbergstr and http://people.clemson.edu/~sbaier).

Political-Stability indicator measures the ability of citizens to peacefully select and replace those in power.

We draw again on the World Bank's governance indicators for variables possibly correlated with negotiation costs between two governments. Negotiation costs may be lower, and the probability of an FTA higher, the larger the capacity of governments to effectively formulate and implement sound policies and the credibility of the government's commitment to such policies. "Government Effectiveness" measures the quality of the bureaucracy, the independence of civil servants, and the credibility of a government's commitments. "Regulatory Quality" focuses on the incidence of "market-unfriendly" policies and perceptions of excessive regulations in foreign trade.

The third theme is the security externality issue. We use an indicator of the "affinity" between two countries for voting similarly at the United Nations' General Assembly over the period 1980-1995, to capture the similarity of two countries in political and military views. This variable may also influence negotiating costs. This measure is described in Gartzke, Jo and Tucker (1999).[19]

3.C. Empirical Results

This section summarizes the results of estimating the gravity equation using data for year 2000 using OLS, IV, and Heckman's control-function approaches. Table 4 presents the results using OLS and IV techniques. Column (1) in Table 4 provides the results of estimating gravity equation (8) above using OLS. As before, OLS suggests little impact of an FTA on bilateral trade.

The remaining specifications in Table 4 use some form of IV technique. For IV, we first need to estimate the probability of an FTA, P(*FTA*). We consider two possibilities, a probit function and a linear probability model. We consider first a probit function. Column (2) in Table 3 reports the results of using IV to estimate the gravity equation, where "identification" is provided by the nonlinearity of the probit function.[20] In this case, we only use Φ^P as an instrument (along with a constant). We assume that Φ^P is determined from a probit estimation of *FTA* on the standard gravity equation variables ($\boldsymbol{q}$); no variables are included in $\boldsymbol{z}$. As shown, the ATE of the FTA is *negative* and statistically significant. Wooldridge himself suggests that – even though feasible – one can "rarely justify" this procedure; for one reason, Φ^P and $\boldsymbol{q}$ are typically highly correlated. Table 5 reports the first-stage regression results for this result and

[19]We note, in advance, that several – if not all – of these variables may be correlated with the error term in the gravity equation. Consequently, our empirical analysis will be sensitive to this, and will provide tests of overidentifying restrictions to help ensure the quality of the results.

[20]Wooldridge (2002, p. 624) discusses the feasibility of using the nonlinearity in the probit function for identification.

all specifications to follow.

Column (3) in Table 4 reports the results of using IV to estimate the gravity equation as described explicitly in section 3A using a probit function. Following Baier and Bergstrand (2004), we estimate a probit function of P(*FTA*) where, motivated by the theoretical model in that paper, economic determinants of the likelihood of a country pair having an FTA include the standard gravity equation variables in ***q*** and a set of instruments ***z***: remoteness of the country pair from the ROW (*REMOTE*), difference in their capital-labor ratios (*DKL*), and the difference in the pair's K-L ratio from that of the ROW (*DROWKL*). In this case, identification of the gravity equation arises from the nonlinearity of the probit function *and* instruments in ***z*** (variables not already in the gravity equation). The results in Table 5 suggest that the instruments identifying the gravity equation coefficients are "good," in the sense of having statistically significant effects in the estimated probit equation. However, two problems remain. First, even though these variables have often been excluded from the typical gravity equation such as (1), in *some* cases these types of variables – differences in K-L ratios and various atheoretical measures of remoteness – have been included in gravity equations and their coefficient estimates have been significant. It is quite feasible that these instruments are correlated with the gravity equation's error term. Second, we can conduct a test of overidentification to determine if the instruments are exogenous. In the case of specification (3), we reject the null hypothesis of exogeneity of the ***z*** instruments [$\chi^2 = 60.61$; $\chi^2_{Crit}(2; 0.01)=9.21$].[21]

While *REMOTE, DKL,* and *DROWKL* are correlated with *FTA*, these "economic" variables are likely correlated with the gravity error term, ϵ_{ij}. This is the major dilemma facing consistent estimation of the gravity equation with IV. If, as Baier and Bergstrand (2004) argue, FTAs have been formed largely to improve economic welfare, any variable likely to be correlated with *FTA* is likely also to be "trade-creating" and thus correlated with ϵ_{ij}. Hence, one needs to consider potential instruments that might influence the decision to form an FTA, but be uncorrelated with potential trade creation. A possibility is a set of *political* variables that might explain *FTA*.

In column (4), we provide the results of estimating the gravity equation using IV where the ***z*** variables in the first-stage are the "political variables." We include the set of six political variables that we consider as likely influencing trade-policy setting and less likely to be related to trade flows, as described in the previous section. However, it remains an empirical matter as to whether these political

[21] Baier and Bergstrand (2002) conducted only limited tests of overidentification and Magee (2003) provided no such tests.

variables are good instruments, as some of these variables have been included as trade determinants in the political science literature. In specification (4), we include in z four measures from the World Bank's governance indicators, a measure of democracy, and an index of affinity of U.N. voting records. In column (4), we find that the use of these "political" instruments does not change earlier results. The ATE for FTA is still negative and statistically insignificant. The χ^2 statistic indicates again that the exogeneity of the instruments can be rejected.[22]

The remaining specifications in Table 4 report the results of using IV, but using a linear probability model (LPM) in the first stage rather than a probit function. There are two reasons for pursuing this strategy. First, as noted in Angrist and Krueger (2001, p. 80), in two-stage least squares (2SLS) consistency of the second-stage estimates does not turn on having the first-stage functional form (usually probit) correct. In fact, using probit may even "do some harm." The second reason for pursuing an LPM first stage is that, with a probit function in the first stage, we cannot use fixed effects in either the first or second stages (to allow for multilateral resistance terms suggested by theory). Using a LPM allows the application later of fixed effects.

Column (5) in Table 4 reports the results of using IV with the first-stage LPM model with economic variables in z, excluding fixed effects. In this case, the ATE of FTA is 2.51, suggesting that an FTA increases bilateral trade *1100 percent*. However, as before the test for overidentifying restrictions suggests that we need to reject exogeneity of the instruments. Column (6) reports the results using political z variables in the first-stage LPM model without fixed effects. The ATE for FTA (2.12) implies a 700 percent effect of an FTA. However, as with the previous regression, one can reject the exogeneity assumption for the instruments using the χ^2 test for overidentifying restrictions. Thus, even though we manage to estimate (perhaps implausibly) large ATEs in the last two specifications. The test for overidentifying restrictions suggests that the results are not reliable.

The final two specifications in Table 4 use a LPM in the first stage and fixed effects in all stages. The purpose of using fixed effects is to allow for any unobservable country-specific (fixed) effects influencing trade flows, such as the multilateral price terms discussed in section 1 for equation (2). To account for endogenous nominal GDPs, we first scale nominal trade flows by the product of the country pair's GDPs, henceforth termed "trade shares."

Columns (7) and (8) report the results for these specifications. First, column (7) reports the

---

[22] We found similar results using various subsets of these political instruments as well as numerous other potential instruments (an exhaustive listing of which we omit for brevity). The results provided here are representative.

results using economic variables *REMOTE, DKL,* and *DROWKL* for ***z***. As shown, the ATE (0.41) is 50 percent larger than the ATE using OLS (0.29), though the ATE is statistically insignificant due to a large coefficient estimation error. Yet, interestingly the test for overidentification suggests we cannot reject exogeneity of the instruments. Column (8) reports the results using the political variables instead for ***z***. The ATE in this case is a large *negative* value. Again, the χ^2 test for overidentifying restrictions implies we cannot reject exogeneity of the instruments. However, the instability of the estimated ATEs is cause for concern.

We conclude that, using a wide (though not exhaustive) array of feasible instrumental variables, IV estimation is not a reliable method for addressing the endogeneity bias of the FTA binary variable in a gravity equation. As addressed earlier, an alternative method for estimating the ATE of FTAs uses Heckman's control-function approach. We have estimated similar specifications using this alternative approach with qualitatively similar findings; the control-function approach does not solve the endogeneity bias issue either. The results are provided in Appendix 3 (Table A1); specification (1) uses the economic ***z*** variables and specification (2) uses the political ***z*** variables.

We have tried a wide set of *potentially* exogenous instrumental variables and Heckman control functions to help identify the gravity equations' coefficient variables with alternative specifications, but with little success. The likely problem is that the vast number of variables that are correlated cross-sectionally with the probability of having an FTA are *also* correlated with trade flows, preventing elimination of the endogeneity bias using cross-section data. The results in this section are largely consistent with the findings in Magee (2003). Magee also estimates using 2SLS the effect of endogenous FTAs on trade flows, and finds similarly a range of large positive to large negative effects of FTAs on trade flows.[23]

However, he concludes that "we should be cautious in using gravity equation estimates to draw strong conclusions about the effect of PTA formation on trade." We agree with his conclusion *for cross-sectional data.* However, in the remainder of this paper, we argue that one *can* draw strong and reliable inferences about the ATE of FTAs using the gravity equation applied to *panel data*.

4. FTA Treatment Effects using Panel Data

As most standard econometrics textbooks now suggest, a ready alternative to cross-section

[23] However, we believe that his *simultaneous* estimation of the effect of bilateral trade flows on influencing the probability of an FTA is flawed because of a logical inconsistency in the underlying structural system of equations. See Appendix 2 and Maddala (1983, p. 118).

estimation of treatment effects in the presence of unobserved heterogeneity is the use of panel data. Having constructed a panel (for every five years) from 1960-2000 of the bilateral trade flows, bilateral trade agreements, and standard gravity equation covariates among 96 potential trading partners, we now pursue this approach.[24] Three main alternative techniques exist for addressing the issue: estimating the panel with fixed effects, estimation with random effects, or differencing the data and using OLS. In section A, we first address fixed versus random effects. In section B, we estimate the model using fixed effects, but ignoring multilateral price terms. In section C, we estimate the model using fixed effects, but accounting for theoretically-motivated multilateral price terms. In section D, we discuss some limitations of estimating the gravity model with fixed effects versus estimation using first-differenced panel data. In section E, we provide results using differenced data, but ignoring theoretically-motivated multilateral price terms. Finally, in section F we estimate the model with differenced data, but accounting for multilateral price terms.

4.A. Fixed versus Random Effects

Our panel estimation applies fixed effects rather than random effects for two reasons, the first on conceptual economic grounds and the second on econometric grounds. First, as addressed in section 2, we believe the source of endogeneity bias in the gravity equation is unobserved heterogeneity. In economic terms, we believe there are unobserved time-invariant bilateral random variables – termed w_{ij} – influencing simultaneously the presence of an FTA and the volume of trade. Even though these variables are random, they are best controlled for using bilateral "fixed effects," as this approach allows for arbitrary correlations of w_{ij} with FTA_{ij}. By contrast, under "random effects" one assumes zero correlation between unobservables w_{ij} with FTA_{ij}, which seems less plausible.

Second, there have been recent econometric evaluations of the gravity equation with panel data using the Hausman Test to test for fixed versus random effects. For example, Egger (2000) finds overwhelming evidence for the rejection of a random-effects gravity model relative to a fixed-effects gravity model, using either bilateral-pair or country-specific fixed effects.

4.B. Fixed-Effects Estimation Ignoring Multilateral Price Terms

In a panel context, equation (1) can be expressed as:

[24] As mentioned earlier, for the panel nominal GDPs are deflated by GDP deflators to generate real GDPs and nominal bilateral trade flows are deflated by the exporter's GDP deflator to generate real trade flows.

$$\ln X_{ijt} = \beta_0 + \beta_1 \ln RGDP_{it} + \beta_2 \ln RGDP_{jt} + \beta_3 \ln DIST_{ij} + \beta_4 (ADJ_{ij}) + \beta_5 (LANG_{ij}) + \beta_6 (FTA_{ijt}) + \varepsilon_{ijt} \quad (9)$$

Table 6 provides the empirical results of estimating gravity equation (9) using a panel of real trade flows (X_{ijt}) and real GDPs ($RGDP_{it}$, $RGDP_{jt}$), and using alternative specifications with and without fixed effects and time dummies. Column (1) provides the baseline gravity equation without any fixed effects or time dummies for all nine years. Exporter and importer (real) GDPs have coefficients close to unity, distance has a traditional coefficient estimate of -1, and the adjacency and language dummies have typical coefficient estimates. The ATE for *FTA* of 0.13 is economically small, consistent with earlier findings in cross-section. Column (2) provides the empirical results including a time dummy, where (for brevity) we omit reporting the (statistically significant) coefficient estimates for these dummy variables. The inclusion of time dummies increases the ATE for *FTA* slightly from 0.13 to 0.27. However, time dummies do not adjust for the endogeneity of FTAs.

Adjusting for unobserved heterogeneity using bilateral fixed effects has a dramatic impact on the results. Column (3) provides results including bilateral fixed effects. The ATE for *FTA* (0.51) is now almost *quadruple* the coefficient estimated using OLS (0.13). Column (4) provides results using both bilateral fixed effects and year dummies. In this specification, the *ATE* of a free trade agreement is 0.68. This quantitative estimate suggests that – accounting for endogeneity bias using fixed effects – the average treatment effect of the presence of a free trade agreement is to *double* trade between country pairs ($e^{0.68}$ = 1.97, or 97 percent increase). This is *seven times* the effect estimated using OLS ($e^{0.13}$ = 1.14, or 14 percent increase).[25]

4.C. Fixed-Effects Estimation Accounting for Multilateral Price Terms

While the results in the previous section are encouraging, the gravity equation suggested by recent formal theoretical developments – summarized in the system of equations (2), (3.1), ..., (3.N) in section 1 – suggests that one needs to account for the *multilateral price variables*. None of the first four specifications in Table 6 accounts for these terms. In a panel context, equation (2) can be expressed as:

[25] The only other published study that has estimated the ATE of an FTA using a panel of data spanning as many years and at least 100 countries is Rose (2004). Using fixed effects, Rose found an ATE of $e^{0.94}$ or 156 percent. However, using a classification of formal and informal GATT members, Tomz (2004) estimates an ATE for FTAs (with fixed effects) of only $e^{0.76}$ or 114 percent, much closer to our results (though still larger due to a larger share of FTAs among small countries in his sample).

$$\ln [X_{ijt} / (GDP_{it}\ GDP_{jt})] = \beta_0 + \beta_3 \ln DIST_{ij} + \beta_4\ (ADJ_{ij})$$
$$+ \beta_5\ (LANG_{ij}) + \beta_6\ (FTA_{ijt}) - \ln P_{it}^{1-\sigma} - \ln P_{jt}^{1-\sigma} + \varepsilon_{ijt} \quad (10)$$

In a panel setting, the multilateral price variables would be *time varying*, and consequently the results in specifications (1)-(4) in Table 6 likely suffer from an omitted variables bias as a result of ignoring these time-varying terms – a dilemma that cannot be resolved by the use of *fixed* effects using the panel data in its current form (levels).[26] Moreover, the theoretical model in equations (2), (3.1), ..., (3.N) suggests that the coefficient estimates for the real GDP variables should be unity. Using bilateral fixed effects in specifications (3) and (4), income elasticities are *significantly* different from unity.

Nevertheless, we can impose the (statistically-rejected) restriction of unitary income elasticities implied by equation (10) and estimate (10) using bilateral-pair (*ij*) fixed effects *along with* country-and-time (*it, jt*) effects. In the context of the theory motivating equation (10), this should generate an unbiased estimate of β_6. Column (5) in Table 6 provides the results of estimating this equation using bilateral fixed effects and country-and-time effects. As indicated, the *FTA* coefficient estimate of 0.46 suggests that an FTA increases trade by about 58 percent on average.

4.D. Fixed-Effects versus First-Differenced Panel Gravity Equation Estimates

Standard econometric discussions of treating endogeneity bias using panel data focus on a choice between estimation using fixed effects versus using first-differenced data, cf., Wooldridge (2002, Ch. 10). As Wooldridge notes, when the number of time periods (T) exceeds two, the fixed-effects estimator is more efficient under the assumption of serially uncorrelated error terms. The first-differencing estimator is more efficient (when $T > 2$) under the assumption that the error term ϵ_{ijt} follows a random walk (i.e., that the difference in the error terms, ϵ_{ijt}-$\epsilon_{ij,t-1}$, is white noise).[27]

First-differencing the panel data yields some potential advantages over fixed effects. First, it is quite plausible that the unobserved heterogeneity in trade flows, ϵ_{ijt}, is correlated over time, in light of discussion in section 2. If the ϵ_{ijt} are highly serially correlated, the inefficiency of fixed effects is exacerbated as T gets large. This suggests that differencing the data will increase estimation efficiency

---

[26] Random effects estimation would not be of any use either, as theory suggests that the multilateral price terms and the FTA variable would be correlated.

[27] When the number of time periods is limited to two (T=2), estimation with fixed effects and first-differencing produce *identical* estimates and inferences; moreover, first-differencing is easier. When T>2, the choice depends upon the assumption the researcher makes about the error term ϵ_{ijt}.

for our large-*T* panel. Second, aggregate trade flow data and real GDP data are likely "close to" unit-root processes. Using fixed effects is equivalent to differencing data around the *mean* (in our sample, 1980); this may create a problem since *T* is large in our panel. As Wooldridge (2000, p. 447) notes, if the data follow unit-root processes and *T* is large, the "spurious regression problem" can arise in a panel using fixed effects. First-differencing yields data that deviates from the previous year of our panel, and thus is closer to a unit-root process. Third, using differenced data may generate income elasticities not significantly different from unity, as suggested by theory. Fourth, first-differenced data will allow consideration of the "phasing-in" of FTAs, as well as allow ready testing of the "strict exogeneity" assumption, as will be discussed.

4.E. First-Differenced Estimation Ignoring Multilateral Price Terms

Table 7 presents the results using alternative specifications with the first-differenced estimator, but *ignoring* the potential omitted variables bias created by the theoretical multilateral price terms. Column (1) presents the basic model. As shown, the ATE for *FTA* is 0.29. Adding time dummies in column (2) has little material effect on the estimated ATE. The smaller FTA effect is not surprising. The reason is that, by differencing the data first, we are evaluating the effect on only a *five-year period* of trade of the formation of an FTA. Since fixed-effects estimation differences data relative to the mean (here, 1980), such estimation potentially measures the effect on a 20-year period of trade of an FTA change.

However, with first-differenced data, one can examine explicitly the effects of *lagged* changes in *FTA* on trade flow changes. The economic motivation for including lagged changes stems from the institutional nature of virtually all FTAs. The 0-1 *FTA* variable was constructed using the "Date of Entry into Force" of the agreement, as best surmised by scrutinizing multiple data sources provided earlier. However, virtually every FTA has "phase-in" periods, generally up to 10 years. For instance, the original EEC agreement of 1958 had a 10-year phase-in period. NAFTA had a similar provision. Thus, the entire economic (treatment) effect cannot be captured fully in the concurrent 5-year period exclusively. It is reasonable to expect an FTA "entered into legally" in a certain 5-year period (e.g., 1960-1965) to not fully come into effect until the next 5-year period (1965-1970). Thus, it is reasonable to include the lagged change in *FTA*.

Moreover, economic effects of an FTA include altering the terms of trade. However, as is well known from a large literature in international economics, terms-of-trade changes tend to have lagged effects on trade volumes. Thus, it is reasonable to assume that a change in *FTA* in period 1960-1965, which is fully "phased-in" by period 1965-1970, might still have an effect on trade flows in period 1970-1975. To allow for these potential impacts, we re-estimate the model in column (2) adding either one lag (column 3) or two lags (column 4).

The results in columns (3) and (4) reveal that changes in *FTA* have lagged effects on trade flow changes. Moreover, the coefficient estimates have economically plausible values. The quantitative similarity of the effects on trade of concurrent and one-period-lagged changes seems consistent with the "phase-ins" of most FTAs; half of the trade effect occurs during the first half of the typical 10-year phase in. Also, second lags have smaller impacts, as is reasonable. The cumulative ATE with one lag is 0.49; with two lags, the total ATE is 0.65. Note how close this estimate (0.65) is to the comparable fixed-effects estimate in column (4) of Table 6 (0.68); the results are robust. The economic interpretation of an ATE of 0.65 is that – after 15 years – the formation of an FTA increases trade by *92 percent*![28]

The introductory quote in this paper questioned whether traditional cross-section estimates of *FTA* on trade can be interpreted as FTAs causing trade *or* trade causing FTAs. The results of previous sections and this section suggest – after accounting for endogeneity using panel data – that one can find economically significant ATEs for *FTA*. However, to confirm that there are no "feedback effects" from trade changes to *FTA* changes, we run one more regression of this model. Wooldridge (2002, p. 285) suggests that it is easy to test for the "strict exogeneity" of FTAs in our context. To do this, we add a *future* change in *FTA* to the regression model. In the panel context here, if *FTA* changes are strictly exogenous to trade flow changes, a future change in *FTA* should be uncorrelated with the concurrent trade flow change. The results in column (4) confirm this. The effect of a future change in *FTA* on the trade flow change is economically small and not significantly different from zero. Moreover, the inclusion of this future change has little material effect on the total ATE of *FTA*. Our results indicate that current innovations in trade are uncorrelated with future decisions to form an FTA.

Also, one can show that the coefficient estimates for the income variables are not significantly different from unity in columns (3) and (4). Thus, real GDPs coefficient estimates are more in line with theoretically-motivated gravity equations.

However, one issue *still* remains. The coefficient estimates in Table 7 may be biased by ignoring

---

[28] Additional lags of *FTA* were also evaluated. None of these extra lags were economically or statistically significant. This suggests that FTAs have a *level* effect on trade, but do not influence the long-run growth rate of trade. Bayoumi and Eichengreen (1997) used first-differenced data to estimate the effects on member trade of the six-member EEC and seven-member EFTA, but did not account for theoretically-motivated time-varying multilateral price resistance terms and only examined one 15-year period for these two agreements (1957-1972). Yet, the implied cumulative ATE for the EEC was 0.60, very similar to our estimated ATE.

the potential omitted variable bias caused by theoretically-motivated time-varying multilateral price terms discussed in section 1. In the next section, we address this issue.

4.F. First-Differenced Estimation Accounting for Multilateral Price Terms

In the context of first-differenced panel data, the potential omitted variables bias created by time-varying multilateral price terms for each country would require country-and-time fixed effects to obtain consistent estimates of *FTA*'s ATE. As earlier, to ensure the theoretically-motivated unit income elasticities in equation (2), we scale the (real) bilateral trade flow from i to j by the product of real GDPs of countries i and j, termed $TradeShare_{ijt}$.[29] Second, we first-difference the natural logarithm of $TradeShare_{ijt}$, creating $dlnTradeShare_{ij,t-(t-1)}$. Third, we regress $dlnTradeShare_{ij,t-(t-1)}$ on 768 country-and-time fixed effects ($Dum_{i,t-(t-1)}$, where *i* denotes a country and *t-(t-1)* a 5-year period, e.g., 1995-2000) and retain the residuals. Fourth, we difference FTA_{ijt}, creating $dFTA_{ij,t-(t-1)}$, and regress $dFTA_{ij,t-(t-1)}$ on the same 768 country-and-time fixed effects and retain these residuals. A regression of the residuals from the first (*dlnTradeShare*) regression on the residuals from the second (*dFTA*) regression will yield unbiased estimates of the ATE effect of an FTA (holding constant time-varying multilateral price terms).

Formally, we estimate:

$$dlnTradeShare_{ij,\ t-(t-1)} = \beta_6\ (dFTA_{ij,\ t-(t-1)}) - dlnP^{1-\sigma}_{i\ ,\ t-(t-1)} - dlnP^{1-\sigma}_{j\ ,\ t-(t-1)} + v_{ij,t-(t-1)} \tag{11}$$

where $v_{ij,t-(t-1)} = \epsilon_{ijt} - \epsilon_{ij,t-1}$ is white noise. With nine years in the panel, we have 8 time periods *t-(t-1)*. Since there are 96 countries that can potentially trade, our procedure above effectively introduces 768 (= 8 x 96) country-and-time fixed effects ($Dum_{i,t-(t-1)}$ and $Dum_{j,t-(t-1)}$) to account for variation in the terms $dlnP^{1-\sigma}_{i,t-(t-1)}$ and $dlnP^{1-\sigma}_{j,t-(t-1)}$ to obtain an unbiased estimate of β_6.

Table 8 reports the coefficient estimates for concurrent, lagged and future changes in *dFTA* on trade flow changes. Column (1) reports the ATE without any lagged effects; the ATE is 0.30. Including one lagged change in *FTA* increases the cumulative ATE to 0.52. With two lagged values of *dFTA*, the cumulative ATE of an FTA is 0.62. This estimate implies that an unbiased ATE of an FTA on the trade between a pair of countries is *86 percent*. As in the previous section, an FTA virtually *doubles* two countries' international trade. This result is larger than the comparable fixed-effects estimate in column (5) of Table 6 (0.46). For completeness, column (4) of Table 8 provides the results of adding a future

[29] Given the results of section 4E, this restriction cannot be rejected.

change in *FTA*. As before, future changes in *FTA* have no significant effect on trade flow changes. Thus, even accounting for time-varying multilateral price terms, our estimates suggest that – after 15 years – an FTA essentially *doubles* members' international trade.

5. Conclusions

The purpose of this paper was to answer the question posed in the title: Do free trade agreements increase members' international trade? A motivation for this paper was that – after forty years of gravity equation estimates of the (treatment) effect of FTAs on trade flows – there seems no clear and convincing empirical evidence using the workhorse for empirical international trade studies that the answer is "yes." This seems surprising in light of the proliferation of FTAs in the last 15 years and widespread expectations that such agreements should increase trade. Our goal has been to provide a thorough empirical analysis of the average treatment effect of FTAs on trade, in light of prevailing knowledge on the theoretical foundations for the trade gravity equation and on modern econometric techniques (developed in the context of labor econometrics) to estimate "average treatment effects."

Our answer to the question is: *yes!* Standard cross-section techniques using instrumental variables and control functions do not provide stable estimates of these ATEs in the presence of endogeneity and overidentifying restrictions tests generally fail. However, we find convincing empirical evidence using panel data of unbiased estimates of ATEs ranging from 0.46 to 0.68. Our preferred ATE estimate from Table 8, column (3) is 0.62. Stated succinctly, this estimate suggests that an FTA will on average increase two member countries' trade about *86 percent* after 15 years (using $e^{0.62} = 1.86$). This is *six times* the effect estimated using OLS ($e^{0.13} = 1.14$, or 14 percent increase).

Some caveats are necessarily in order. While we have addressed general equilibrium impacts of an FTA on the multilateral price (resistance) terms of a country pair, we have not addressed the impact of such agreements on trade with nonmembers, nor trade among nonmembers. Moreover, we have not addressed the welfare implications of FTAs. These are topics left for other research; our focus has been solely on trying to provide policymakers with more resolution on an unbiased estimate of the average effect of an FTA on trade. A second caveat is that the effect of an FTA is likely to differ depending upon the agreement. We consider this a useful extension for future research. A related caveat to the previous one is that the effect of an FTA on two members' trade is likely influenced by the economic size, per capita incomes, and even distance between the two countries. That is, there are likely interactive effects not addressed in this study, to limit its scope. This topic is left for future research.

Nevertheless, in light of the wide range of previous estimates of the effect of FTAs on trade, we hope that future research will acknowledge the importance of addressing the endogeneity of FTAs, as the international trade literature recognized years ago for multilateral trade volumes and countries' tariff and nontariff barrier policies. We hope that future research will come to appreciate the importance of adjusting for endogeneity when estimating the effects of free trade agreements on trade flows.

Table 1

Typical Gravity Equation Coefficient Estimates
(dependent variable is natural log of bilateral trade flow from i to j)

Variable	(1) 1960	(2) 1970	(3) 1980	(4) 1990	(5) 2000
ln GDP_i	0.76	0.88	1.01	1.08	1.18
	(45.79)	(57.55)	(69.37)	(85.13)	(104.13)
ln GDP_j	0.76	0.92	1.00	0.97	0.98
	(48.66)	(63.95)	(72.69)	(78.08)	(87.39)
ln $DIST_{ij}$	-0.64	-0.85	-1.06	-1.07	-1.17
	(-16.23)	(-21.10)	(-28.15)	(-28.82)	(-32.57)
ADJ_{ij}	0.16	0.14	0.35	0.59	0.74
	(1.03)	(0.85)	(2.18)	(3.72)	(4.88)
$LANG_{ij}$	0.06	0.34	0.56	0.80	0.72
	(0.65)	(3.48)	(5.84)	(8.16)	(7.71)
FTA_{ij}	0.63	1.37	-0.13	-0.14	0.29
	(3.46)	(6.64)	(-0.73)	(-0.95)	(2.85)
Constant	-9.38	-12.17	-16.23	-17.09	-17.94
	(-20.44)	(-26.88)	(-35.59)	(-40.37)	(-49.11)
RMSE	1.4163	1.7616	1.8900	1.9919	1.9645
R^2	0.6061	0.6334	0.6446	0.6649	0.7137
No. observ.	2633	4030	5421	6474	7302

Table 2

Theory-Motivated Gravity Equation with Country Fixed Effects
(dependent variable is natural log of bilateral trade flow divided by product of nominal GDPs)

Variable	(1) 1960	(2) 1970	(3) 1980	(4) 1990	(5) 2000
ln $DIST_{ij}$	-0.68	-0.89	-1.28	-1.30	-1.46
	(-16.77)	(-21.58)	(-31.36)	(-31.65)	(-35.79)
ADJ_{ij}	0.31	0.35	0.43	0.58	0.59
	(2.26)	(2.38)	(2.95)	(3.93)	(4.09)
$LANG_{ij}$	0.38	0.84	0.82	0.98	0.97
	(3.99)	(8.33)	(8.06)	(9.41)	(9.78)
FTA_{ij}	0.01	0.61	-1.44	-1.08	-0.14
	(0.09)	(3.27)	(-8.65)	(-7.30)	(-1.36)
Constant	-14.06	-12.49	-14.98	-16.64	-12.76
	(-8.25)	(-18.66)	(-19.37)	(-31.88)	(-27.18)
RMSE	1.1826	1.5025	1.6635	1.7806	1.7851
R^2	0.5020	0.4300	0.3857	0.3648	0.3845
No. observ.	2633	4030	5421	6474	7302

Table 3

Free Trade Agreements
(in chronological order of date of entry into force)

European Union, or EU (1958): Austria (1995), Belgium-Luxembourg, Denmark (1973), Finland (1995), France, Greece (1981), Ireland (1973), Italy, Germany, Netherlands, Portugal (1986), Spain (1986), Sweden (1995), United Kingdom (1973).

The Customs Union of West African States (1959): Burkina Faso, Mali, Mauritania, Niger, Senegal.

European Free Trade Association, or EFTA (1960): Austria (until 1995), Denmark (until 1973), Finland (1986-1995), Norway, Portugal (until 1986), Sweden (until 1995), Switzerland, United Kingdom (until 1973).

Latin American Free Trade Agreement/Latin American Integration Agreement, or LAFTA/LAIA (1961-1979,1993-): Argentina, Bolivia, Brazil, Chile, Ecuador, Mexico, Paraguay, Peru, Uruguay, Venezuela (ineffective 1980-1990 reinitiated 1993).

African Common Market (1963): Algeria, Egypt, Ghana, Morocco.

Central American Common Market (1961-1975, 1993-): El Salvador, Guatemala, Honduras, Nicaragua, Costa Rica (1965).

Economic Customs Union of the Central African States (1966): Cameroon, Congo, Gabon.

Carribean Community, or CARICOM (1968): Jamaica, Trinidad and Tobago, Guyana (1995).

EU-EFTA/European Economic Area (1973/1994)

Australia -New Zealand Closer Economic Relations (1983).

US-Israel (1985)

US-Canada (1989)

EFTA-Israel (1993)

Central Europe Free Trade Agreement, or CEFTA (1993): Bulgaria (1998), Hungary, Poland, Romania (1997).

EFTA-Bulgaria (1993)

EFTA-Hungary (1993)

EFTA-Poland (1993)

EFTA-Romania (1993)

EU-Hungary (1994)

EU-Poland (1994)

NAFTA (1994): Canada, Mexico, United States.

Bolivia-Mexico (1995)

Carribean Community - Dominican Republic (1995)

Costa Rica-Mexico (1995)

EU-Bulgaria (1995)

EU-Romania (1995)

Group of Three (1995): Columbia, Mexico, Venezuela

Mercado Comun del Sur, or Mercosur (1991): Argentina, Brazil, Paraguay, Uruguay (formed in 1991 and a free trade area in 1995).

Carribean Community-Dominican Republic (1998)

Common Market for Eastern and Southern Africa (effective 1998): Egypt, Kenya, Madagascar, Malawi, Mauritius, Sudan, Zimbabwe, Zambia.

Andean Community (1995): Columbia, Ecuador, Peru Venezuela. Customs Union: Columbia, Ecuador, Venezuela. Free Trade Area: Bolivia (1995), Peru (1997).

Mercosur-Chile (1996)

Mercosur-Bolivia (1996)

Canada-Chile (1997)

Canada-Israel (1997)

Association of Southeast Nations, or ASEAN (1998): Indonesia, Philippines, Singapore, Thailand. (Effective on 80% merchandise trade in 1998).

Hungary-Turkey (1998)

Hungary-Israel (1998)

India-Sri Lanka (1998)

Israel-Turkey (1998)

Mexico-Nicaragua (1998)

Romania-Turkey (1998)

Poland-Israel (1998)

Romania-Turkey (1998)

Mexico-Chile (1999)

EU-Israel Agreement (2000)

EU-Mexico (2000)

Poland-Turkey (2000)

Mexico-Guatemala (2000)

Mexico-Honduras (2000)

Mexico-Israel (2000)

Mexico-El Salvador (2000)

New Zealand-Singapore (2000)

Sources:

http://www.wto.org/english/tratop_e/region_e/summary_e.xls
http://europa.eu.int/comm/enlargement/pas/europe_agr.htm
http://www.comunidadandina.org/ingles/union.htm
http://www.nafinsa.com/finsafreetrade.htm
http://www.sice.oas.org/default.asp
Frankel, Jeffrey A., Ernesto Stein, She-Jin Wei. *Regional Trading Blocs*. Institute for International Economics (1997).
Lawrence, Robert Z. *Regionalism, Multilateralism, and Deeper Integration.* The Brookings Institution (1996).

Table 4

Gravity Equations using OLS and IV Techniques
(dependent variable is natural log of bilateral trade flow in year 2000)

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln GDP_i	1.18	1.20	1.19	1.18	1.13	1.14		
	(103.23)	(95.13)	(99.96)	(97.01)	(77.15)	(79.94)		
ln GDP_j	0.98	0.99	0.98	0.99	0.93	0.94		
	(89.43)	(79.89)	(83.86)	(81.93)	(63.86)	(66.83)		
ln $DIST_{ij}$	-1.17	-1.30	-1.21	-1.26	-0.85	-0.97	-1.37	-2.11
	(-32.50)	(-24.93)	(-27.85)	(-28.22)	(-11.75)	(-14.56)	(-18.41)	(-15.30)
ADJ_{ij}	0.74	0.86	0.78	0.80	0.46	0.52	0.53	1.01
	(4.86)	(5.49)	(5.06)	(4.98)	(2.76)	(3.10)	(3.52)	(5.53)
$LANG_{ij}$	0.72	0.77	0.73	0.76	0.58	0.62	0.97	1.07
	(8.30)	(8.16)	(7.84)	(7.99)	(5.83)	(6.18)	(9.70)	(9.59)
FTA_{ij}	0.29	-0.62	0.01	-0.03	2.51	2.12	0.41	-3.97
	(3.36)	(-2.20)	(0.07)	(-0.15)	(5.55)	(5.06)	(1.03)	(-4.84)
Constant	-19.69	-19.22	-19.54	-19.13	-20.83	-20.18	-14.64	-9.27
	(-47.79)	(-47.07)	(-49.71)	(-47.46)	(-45.80)	(-45.20)	(-20.87)	(-8.34)
RMSE	1.9645	1.9751	1.9655	1.9771	2.0254	2.0142	1.7886	1.9617
R^2	0.7137	0.7106	0.7134	0.7104	0.6957	0.6994	0.3821	0.2697
No. observ.	7302	7302	7302	6892	7302	6892	7302	6892
χ^2	NA	NA	60.61	263.27	18.99	217.79	2.19	0.00

Table 5

First-Stage Probit and Linear Probability Model Estimates to Generate Predicted P(*FTA*) Associated with Table 3 Results

Variable	(1)	(2)[a]	(3)[a]	(4)[a]	(5)[b]	(6)[b]	(7)[b]	(8)[b]
Basic Gravity Variables:								
$\ln GDP_i + \ln GDP_j$		0.18	0.17	0.13	0.02	0.01		
		(12.34)	(10.06)	(6.72)	(17.42)	(9.80)		
$\ln DIST_{ij}$		-1.10	-0.80	-1.06	-0.09	-0.12	-0.09	-0.13
		(-27.95)	(-16.81)	(-25.06)	(-24.45)	(-38.36)	(-20.94)	(-35.13)
ADJ_{ij}		-0.14	-0.24	-0.02	0.12	0.13	0.12	0.11
		(-1.25)	(-2.03)	(-0.18)	(8.03)	(8.41)	(8.73)	(7.52)
$LANG_{ij}$		0.55	0.54	0.30	0.06	0.04	0.04	0.02
		(6.58)	(5.92)	(3.11)	(6.50)	(3.86)	(3.70)	(1.77)
Economic Instruments:								
$REMOTE_{ij}$			0.09		0.01		0.01	
			(9.34)		(11.68)		(13.90)	
DKL_{ij}			-0.52		-0.02		0.01	
			(-10.95)		(-11.50)		(1.70)	
$DROWKL_{ij}$			0.15		0.01		0.09	
			(3.17)		(4.19)		(11.79)	
Political Instruments:								
ACCOUNT-$ABILITY_{ij}$				0.46		0.04		0.02
				(6.78)		(9.26)		(1.95)
$STABILITY_{ij}$				-0.31		-0.02		-0.04
				(-5.31)		(-4.50)		(-5.03)
$EFFECTIVENESS_{ij}$				-0.09		-0.01		0.06
				(-1.44)		(-1.36)		(5.22)

Table 5 (cont.)

Variable	(1)	(2)[a]	(3)[a]	(4)[a]	(5)[b]	(6)[b]	(7)[b]	(8)[b]
$REGULATORY_{ij}$				0.39		0.03		-0.00
				(4.88)		(5.82)		(-0.43)
$DEMOCRACY_{ij}$				-0.23		0.01		0.00
				(-1.92)		(1.59)		(0.17)
U.N. $AFFINITY_{ij}$				1.02		0.12		0.16
				(7.64)		(12.92)		(13.70)
Constant		0.54	-1.51	1.10	0.20	0.56	0.87	0.91
		(1.37)	(-2.99)	(1.99)	(4.70)	(13.57)	(17.81)	(19.99)
R^2					0.2809	0.2766	0.4094	0.3601
Pseudo R^2		0.4347	0.5123	0.4792				
No. observ.		9120	9120	8702	9120	8702	9120	8702

[a]z-statistics in parentheses.
[b]t-statistics in parentheses.

Table 6

Panel Gravity Equation Estimates in Levels[a]
(dependent variable is natural log of real bilateral trade flows)

Variable	(1) No Fixed or Time Effects	(2) With Time Effects	(3) With Bilateral Fixed Effects	(4) With Time and Bilateral Fixed Effects	(5) Bilateral Fixed and Country-and-Time Effects
ln $RGDP_i$	0.95	0.97	0.71	1.27	
	(217.50)	(230.98)	(34.54)	(47.16)	
ln $RGDP_j$	0.94	0.97	0.58	1.22	
	(224.99)	(235.43)	(26.57)	(41.60)	
ln $DIST_{ij}$	-1.03	-1.01			
	(-79.09)	(-78.60)			
ADJ_{ij}	0.41	0.38			
	(8.23)	(7.28)			
$LANG_{ij}$	0.63	0.58			
	(19.06)	(17.73)			
FTA_{ij}	0.13	0.27	0.51	0.68	0.46
	(3.73)	(7.19)	(10.74)	(14.27)	(9.07)
RMSE	1.9270	1.8601			
Overall R^2	0.6575	0.6809	0.5787	0.6096	0.0237
Between R^2			0.6340	0.6648	0.0091
Within R^2			0.2036	0.2268	0.1892
No. observ.	47081	47081	47081	47081	47081

[a]t-statistics in parentheses.

Table 7

First-Differenced Panel Gravity Equation Estimates Ignoring Multilateral Price Terms[a]

Variable	(1) No Time Dummies or Lags	(2) With Time Dummies	(3) With Time Dummies and Lags	(4) With Time Dummies and 2 Lags	(5) With Time Dummies, 2 Lags, and 1 Future dFTA
dln $RGDP_i$	0.85**	0.87**	0.89**	1.03**	0.98**
	(16.57)	(16.16)	(15.92)	(16.61)	(18.83)
dln $RGDP_j$	1.08**	1.15**	1.12**	1.11**	1.13**
	(18.86)	(19.09)	(17.93)	(17.04)	(18.16)
$dFTA_{ij}$	0.29**	0.26**	0.25**	0.24**	0.21**
	(8.72)	(7.81)	(7.46)	(7.39)	(2.62)
$(dFTA_{ij})_{-1}$			0.24**	0.24**	0.20*
			(6.47)	(6.24)	(2.10)
$(dFTA_{ij})_{-2}$				0.17**	0.19*
				(2.90)	(1.76)
$(dFTA_{ij})_{+1}$					0.06
					(0.89)
Constant	-0.11**	-0.26**	-0.26**	-0.16**	-0.08**
	(-7.13)	(-9.59)	(-9.52)	(-5.81)	(-2.82)
Total ATE	0.29	0.26	0.49	0.65	0.60
RMSE	1.3234	1.3203	1.3474	1.3755	
Overall R^2	0.0259	0.0307	0.0304	0.0325	0.0364
No. observ.	36563	36563	34105	31172	24662

[a]t-statistics in parentheses. * (**) denotes statistical significance at 5 (1) percent level in one-tailed t-test.

Table 8

First-Differenced Panel Gravity Equation Estimates with Country-and-Time Fixed Effects to Account for Multilateral Price Terms[a]

Variable	(1)	(2)	(3)	(4)
$dFTA_{ij}$	0.30**	0.29**	0.28**	0.29**
	(7.90)	(7.58)	(7.41)	(6.77)
$(dFTA_{ij})_{-1}$		0.23**	0.21**	0.16**
		(5.09)	(4.52)	(2.62)
$(dFTA_{ij})_{-2}$			0.13*	0.09
			(2.12)	(1.12)
$(dFTA_{ij})_{+1}$				-0.06
				(-1.02)
Constant	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Total ATE	0.30	0.52	0.62	0.54
RMSE	1.2767	1.3039	1.3337	1.3160
R^2	0.0007	0.0009	0.0008	0.0006
No. observ.	36563	34105	31172	24642

[a]t-statistics in parentheses. * (**) denotes statistical significance at 5 (1) percent level in one-tailed t-test.

APPENDIX 1: DATA

The following is a list of the 96 countries potentially used in the regressions, depending upon availability of non-zero and non-missing trade flows:

Austria	Belgium-Luxembourg	Denmark
Finland	France	Germany
Greece	Ireland	Italy
Netherlands	Norway	Portugal
Spain	Sweden	Switzerland
United Kingdom	Canada	Costa Rica
Dominican Republic	El Salvador	Guatemala
Haiti	Honduras	Jamaica
Mexico	Nicaragua	Panama
Trinidad and Tobago	United States	Argentina
Bolivia	Brazil	Chile
Colombia	Ecuador	Guyana
Paraguay	Peru	Uruguay
Venezuela	Australia	New Zealand
Bulgaria	Hungary	Poland
Romania	Egypt	India
Japan	Philippines	Thailand
Turkey	Korea	Algeria
Angola	Ghana	Kenya
Morocco	Mozambique	Nigeria
Tunisia	Uganda	Zambia
Zimbabwe	China (Hong Kong)	Indonesia
Iran	Israel	Pakistan
Singapore	Sri Lanka	Syrian Arab Republic
China,P.R.: Mainland	Albania	Bangladesh
Burkina Faso	Cameroon	Cyprus
Côte d'Ivoire	Ethiopia	Gabon
Gambia, The	Guinea-Bissau	Madagascar
Malawi	Malaysia	Mali
Mauritania	Mauritius	Niger
Saudi Arabia	Senegal	Sierra Leone
Sudan	Congo, Dem. Rep. of	Congo, Republic of

APPENDIX 2: SIMULTANEITY

To examine the simultaneity bias between U.S. multilateral trade volumes and nontariff policies, Trefler (1993) and Swagel and Lee (1997) employed 2SLS to cross-section data. In the spirit of this approach, Magee (2003) employed 2SLS to address the potential simultaneity bias between bilateral trade flows and an FTA dummy variables using cross-section data. We have not discussed the possibility of estimating a system of simultaneous equations in x and FTA, as pursued in Magee (2003). Our principal reason stems from a discussion in Maddala (1983, p. 118). A natural inclination is to consider the simple theoretical simultaneous equations system:

$$x = \mu_0 + \alpha\, FTA + \beta'\boldsymbol{q} + u_x \qquad (2.1)$$
$$fta = \xi + \lambda\, x + \delta'\boldsymbol{z} + u_{fta} \qquad (2.2)$$

where x is the trade flow, FTA is a dummy variable, fta is a latent variable representing the utility gain or loss for the representative consumer of an FTA as discussed in Baier and Bergstrand (2004), $\boldsymbol{q}$ are the standard gravity equation RHS variables, and $\boldsymbol{z}$ is a set of instruments (not in $\boldsymbol{q}$). Utility gains or losses from an FTA may be a function of variables in $\boldsymbol{z}$, but also may be a function of x. However, we cannot observe fta, so instead we assume the indicator variable, FTA, is a function of fta, $FTA(fta>0) = 1$ and 0 otherwise, and $P(FTA = 1) = P(fta > 0) = \Phi(\ )$, where $\Phi(\ )$ is the standard normal cumulative distribution function. This is the approach in Magee (2003).

However, as summarized in Maddala (1983, p. 118), for logical consistency either λ or α must equal zero. If $\alpha = 0$, then studying the effect of an FTA on trade is irrelevant. So λ must equal 0 and fta cannot *logically* be a function of x. To see this, substitute equation (2.1) into (2.2) to yield $fta = \xi + \lambda\mu_0 + \lambda\alpha\, FTA + \lambda\beta'\boldsymbol{q} + \delta'\boldsymbol{z} + \lambda u_{spx} + u_{fta}$. If $fta > 0$, then $FTA = 1$ implying $P(FTA = 1) = P(fta > 0) = \Phi[\xi + \lambda\mu_0 + \lambda\alpha + \lambda\beta'\boldsymbol{q} + \delta'\boldsymbol{z} + \lambda u_x]$. If $fta \leq 0$, then $FTA = 0$ implying $P(FTA = 0) = P(fta \leq 0) = 1 - \Phi[\xi + \lambda\mu_0 + \lambda\beta'\boldsymbol{q} + \delta'\boldsymbol{z} + \lambda u_x]$. For logical consistency, $P(FTA = 0) + P(FTA = 1) = 1$. Hence, the model is consistent only if:

$$1 - \Phi[\xi + \lambda\mu_0 + \lambda\beta'\boldsymbol{q} + \delta'\boldsymbol{z} + \lambda u_x] + \Phi[\lambda\alpha + \xi + \lambda\mu_0 + \lambda\beta'\boldsymbol{q} + \delta'\boldsymbol{z} + \lambda u_x] = 1\ .$$

This condition holds *if and only* if either λ or α equals zero (so that $\lambda\alpha = 0$). Thus, estimation of the system in equations (2.1) and (2.2) is logically inconsistent unless λ or α equals zero.

Even if one ignores this logical inconsistency, there are two reasons to suggest that the results in Magee might be biased and inconsistent. First, when using IV, Magee does not use a three-stage procedure as in our study; he calculates the predicted probability of an FTA using a probit, and then inserts this point estimate into the trade flow equation. As discussed in Angrist and Krueger (2001, p. 80), this will yield inconsistent estimates in the final-stage regression *unless* the probit function "happens to be exactly right." In our model, the predicted probabilities are used in a linear regression along with $\boldsymbol{q}$ to generate consistent final-stage estimates. Second, Magee includes in his $\boldsymbol{z}$ variables the average (of the logs) of per capita incomes. However, as is well known, the average of per capita incomes is traditionally an *important* explanatory variable in the gravity equation. Its absence in his gravity model implies that the error term in his gravity equation is likely to be highly correlated with $\boldsymbol{z}$, violating a necessary condition that $\boldsymbol{z}$ must satisfy for IV.

APPENDIX 3: HECKMAN CONTROL FUNCTIONS

Intuitively, think of the sample as decomposable into two subsets, x_1 and x_0. Define $E(x_1)$ and $E(x_0)$ as:

$$E(x_1 \mid \boldsymbol{q}, z, FTA{=}1) = \mu_1 + \beta'\boldsymbol{q} + E(\epsilon_1 \mid \boldsymbol{q}, z, FTA{=}1) \tag{3.1}$$

and

$$E(x_0 \mid \boldsymbol{q}, z, FTA{=}0) = \mu_0 + \beta'\boldsymbol{q} + E(\epsilon_0 \mid \boldsymbol{q}, z, FTA{=}0) \tag{3.2}$$

Then the expectation of equation (5) becomes:

$$E(x \mid \boldsymbol{q}, z) = (FTA)\ E(x_1 \mid \boldsymbol{q}, z, FTA{=}1) + (1\text{-}FTA)\ E(x_0 \mid \boldsymbol{q}, z, FTA{=}0) \tag{3.3}$$

or

$$E(x \mid \boldsymbol{q}, z) = \xi + \alpha\ FTA + \beta'\boldsymbol{q} + FTA\ E(\epsilon_1 \mid \boldsymbol{q}, z, FTA{=}1) + (1\text{-}FTA)\ E(\epsilon_0 \mid \boldsymbol{q}, z, FTA{=}0) \tag{3.4}$$

If we assume ϵ_1 and ϵ_0 are joint normally distributed, then we can write:

$$E(x \mid \boldsymbol{q}, z) = \xi + \alpha\ FTA + \beta'\boldsymbol{q} + \rho\{FTA[\phi(\boldsymbol{q}, z; \pi)/\ \Phi(\boldsymbol{q}, z; \pi)] - (1\text{-}FTA)[-\phi(\boldsymbol{q}, z; \pi)/\{1-\ \Phi(\boldsymbol{q}, z; \pi)\}]\} \tag{3.5}$$

where $\phi(\boldsymbol{q}, z; \pi) = \phi(\pi_0 + \pi_1'\boldsymbol{q} + \pi_2'z)$ is the standard normal probability density function, $\Phi(\boldsymbol{q}, z; \pi) = \Phi(\pi_0 + \pi_1'\boldsymbol{q} + \pi_2'z)$ is the standard normal cumulative distribution function, and ρ is a parameter (positively defined). $\{FTA[\phi(\ )/\Phi(\ )] - (1\text{-}FTA)[-\phi(\ )/\{1-\Phi(\ )\}]\}$ serves as a "control" function for the unobservable heterogeneity and potential selection bias, and is often referred to as the inverse Mills ratio or hazard rate.

This analysis suggests a two-step estimator. First, estimate π_0, π_1, and π_2 from a probit of FTA on $\boldsymbol{q}$, z, and a constant to form predicted probabilities Φ^P along with predicted ϕ^P for all observations. Second, estimate by OLS:

$$x = \xi + \alpha\ FTA + \beta'\boldsymbol{q} + \rho\ \{FTA\ [\phi^P / \Phi^P] - (1\text{-}FTA)[-\phi^P/(1-\ \Phi^P)]\} + u \tag{3.6}$$

where u is a normally distributed error term. We will refer to this estimation technique as the "Heckman" control-function procedure. The coefficients will be consistently estimated; see Heckman (1997), Vella and Verbeek (1999), and Wooldridge (2002, Ch. 18). In Table A1, we provide results from this technique.

Table A1

Gravity-Equation Estimates using Heckman Control Function
(dependent variable is natural log of bilateral trade flow in year 2000)

Variables	(1) With Economic Variables in *z*	(2) With Political Variables in *z*
ln GDP_i	1.19	1.18
	(99.53)	(96.49)
ln GDP_j	0.98	0.98
	(83.52)	(81.53)
ln $DIST_{ij}$	-1.21	-1.24
	(-27.61)	(-27.56)
ADJ_{ij}	0.78	0.78
	(5.07)	(4.85)
$LANG_{ij}$	0.73	0.75
	(7.85)	(7.88)
FTA_{ij}	0.01	0.14
	(0.63)	(0.63)
HAZARD RATE	0.20	0.14
	(1.62)	(1.12)
Constant	-19.55	-19.21
	(-49.76)	(-47.72)
RMSE	1.9643	1.9753
R^2	0.7138	0.7109
No. observ.	7302	6892

REFERENCES

Abrams, R.K. "International Trade Flows under Flexible Exchange Rates." Federal Reserve Bank of Kansas City *Economic Review* 65, no. 3 (1980): 3-10.

Aitken, Norman D. "The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis." *American Economic Review* 5 (December 1973): 881-892.

Anderson, James E. "A Theoretical Foundation for the Gravity Equation." *American Economic Review* 69, no. 1 (March 1979): 106-116.

Anderson, James E., and van Wincoop, Eric. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93, no. 1 (March 2003): 170-192.

Angrist, Joshua D., and Krueger, Alan B. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15, no. 4 (Fall 2001): 69-86.

Baier, Scott L., and Bergstrand, Jeffrey H. "The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity." *Journal of International Economics* 53, no. 1 (February 2001): 1-27.

Baier, Scott L., and Bergstrand, Jeffrey H. "On the Endogeneity of International Trade Flows and Free Trade Agreements." Manuscript, March 2002.

Baier, Scott L., and Bergstrand, Jeffrey H. "Economic Determinants of Free Trade Agreements." *Journal of International Economics* 64, no. 1 (September 2004): 29-63.

Baier, Scott L., Dwyer, Gerald, and Tamura, Robert. "The Growth and Development of Nations." Manuscript, 2003.

Baldwin, Richard E. *Towards and Integrated Europe*. London: Centre for Economic Policy Research, 1994.

Bayoumi, Tamim, and Eichengreen, Barry. "Is Regionalism Simply a Diversion? Evidence from the Evolution of the EC and EFTA." In *Regionalism vs. Multilateral Arrangements*, edited by Takatoshi Ito and Anne O. Krueger. Chicago: The University of Chicago Press, 1997.

Bergstrand, Jeffrey H. "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence." *Review of Economics and Statistics* 67, no. 3 (August 1985): 474-481.

Bergstrand, Jeffrey H. "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade." *Review of Economics and Statistics* 71, no. 1 (February 1989): 143-153.

Bergstrand, Jeffrey H. "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis, and the Determinants of Bilateral Intra-Industry Trade." *Economic Journal* 100, no. 4 (December 1990), 1216-1229.

Boisso, Dale, and Ferrantino, Michael. "Economic Distance, Cultural Distance, and Openness in International Trade: Empirical Puzzles." *Journal of Economic Integration* 12, no. 4 (December 1997): 456-484.

Brada, Josef C., and Mendez, Jose A. "Economic Integration among Developed, Developing and Centrally

Planned Economies: A Comparative Analysis." *Review of Economics and Statistics* 67, no. 4 (November 1985): 549-556.

Cheng, I-Hui, and Wall, Howard J. "Controlling for Heterogeneity in Gravity Models of Trade," Federal Reserve Bank of St. Louis Working Paper No. 1999-010C (2002).

Deardorff, Alan. "Does Gravity Work in a Neoclassical World?" In *The Regionalization of the World Economy*, edited by Jeffrey A. Frankel. Chicago: The University of Chicago Press, 1998

Eaton, Jonathan, and Kortum, Samuel. "Technology, Geography, and Trade." *Econometrica* 70, no. 5 (September 2002): 1741-1779.

Egger, Peter. "A Note on the Proper Econometric Specification of the Gravity Equation," *Economics Letters* 66 (2000): 25-31.

Eichengreen, Barry, and Irwin, Douglas. "Trade Blocs, Currency Blocs and the Reorientation of World Trade in the 1930s." *Journal of International Economics* 38, no. ½, (February 1995): 1-24..

Evenett, Simon, and Keller, Wolfgang. "On Theories Explaining the Success of the Gravity Equation." *Journal of Political Economy* 110, no. 2 (April 2002): 281-316.

Feenstra, Robert. *Advanced International Trade: Theory and Evidence*. Princeton, NJ: Princeton University Press, forthcoming 2003.

Felbermyer, Gabriel J., and Kohler, Wilhelm. "Exploring the Intensive and Extensive Margins in World Trade." Manuscript, August 2004.

Frankel, Jeffrey A. *Regional Trading Blocs*. Washington, DC: Institute for International Economics, 1997.

Frankel, Jeffrey A., Ernesto Stein, and Shang-Jin Wei. "Trading Blocs and the Americas: The Natural, the Unnatural, and the Super-Natural." *Journal of Development Economics* 47, no. 1 (June 1995): 61-95.

Frankel, Jeffrey A., Stein,Ernesto and Wei, Shang-Jin. "Regional Trading Arrangements: Natural or Supernatural?" *American Economic Review* 86, no. 2 (May 1996): 52-56.

Frankel, Jeffrey A., Stein, Ernesto, and Wei, Shang-Jin. "Continental Trading Blocs: Are They Natural or Supernatural?" In *The Regionalization of the World Economy*, edited by Jeffrey A. Frankel. Chicago: The University of Chicago Press, 1998, 91-113.

Gartzke, Eric, Jo, Doong-Joon, and Tucker, Richard. "United Nations General Assembly Voting," August 1999 (http://www.vanderbilt.edu/~rtucker/data/affinity/un/assembly).

Ghosh, Sucharita, and Yamarik, Steven. "Are Regional Trading Arrangements Trade Creating?: An Application of Extreme Bounds Analysis." *Journal of International Economics* 63, no. 2 (July 2004): 369-395.

Gilpin, Robert. *The Challenge of Global Capitalism*. Princeton, NJ: Princeton University Press, 2000.

Goldberg, Pinelopi Koujianou, and Maggi, Giovanni. "Protection for Sale: An Empirical Investigation." *American Economic Review* 89, no. 5 (December 1999): 1135-1155.

Gowa, Joanne, and Mansfield, Edward D. "Power Politics and International Trade." *American Political Science Review* 87 (1993): 408-420.

Grossman, Gene M., and Helpman, Elhanan. "The Politics of Free-Trade Agreements." *American Economic Review* 85, no. 4 (September 1995): 667-690.

Head, Keith, and Ries, John. "Increasing Returns vs. National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade." *American Economic Review* 91, no. 4 (September 2001): 858-876.

Heckman, James J. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46, no. 1 (July 1978): 931-959.

Heckman, James J. "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources* 32, no. 3 (Summer 1997): 441-462.

Heckman, James J. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109, no. 4 (August 2001): 673-748.

Helliwell, John F. *How Much Do National Borders Matter?* Washington, DC: The Brookings Institution, 1998.

Helpman, Elhanan, and Krugman, Paul. *Market Structure and Foreign Trade*. Cambridge, MA: MIT Press, 1985.

Jaggers, Keith, and Gurr, Robert Tedd. "Transitions to Democracy: Tracking Democracy's third Wave with Polity III Data," *Journal of Peace Research* 32 (November 1999), 469-482.

Kaufmann, Daniel, Kraay, Aart, and Mastruzzi, M. "Governance Matters III: Governance Indicators for 1996-2002," *World Bank Policy Research Working Paper* No. 3106 (2003).

Lawrence, Robert Z. *Regionalism, Multilateralism, and Deeper Integration.* Washington, DC: The Brookings Institution, 1996.

Lawrence, Robert Z. "Comment on 'The Role of History in Bilateral Trade Flows'." In *The Regionalization of the World Economy*, edited by Jeffrey A. Frankel. Chicago: The University of Chicago Press, 1998, 57-59.

Lee, Jong-Wha, and Swagel, Phillip. "Trade Barriers and Trade Flows Across Countries and Industries." *Review of Economics and Statistics* 79 (1997): 372- 382.

Linnemann, Hans. *An Econometric Study of International Trade Flows*. Amsterdam: North-Holland, 1966.

Maddala, G.S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, MA: Cambridge University Press, 1983.

Magee, Chris. "Endogenous Preferential Trade Agreements: An Empirical Analysis." *Contributions to Economic Analysis and Policy* 2, no. 1. Berkeley Electronic Press, 2003.

Mansfield, Edward D., Helen V. Milner, and B. Peter Rosendorff. "Why Democracies Cooperate More: Electoral Control and International Trade Agreements." *International Organization* 56 (2002): 477-514.

McCallum, John. "National Borders Matter: Canada-U.S. Regional Trade Patterns." *American Economic Review* 85, no. 3 (June 1995): 615-623.

Mitchell, Brian R. *International Historical Statistics: Europe 1750-1988*. New York: Stockton Press, 1992.

Mitchell, Brian R. *International Historical Statistics: The Americas 1750-1988*. New York: Stockton Press, 1993.

Mitchell, Brian R. *International Historical Statistics: Africa, Asia and Oceania 1750-1988*. New York: Stockton Press, 1995.

Oguledo, Victor I., and MacPhee, Craig R. "Gravity Models: A Reformulation and an Application to Discriminatory Arrangements." *Applied Economics* 26 (1994): 107-120.

Preeg, Ernest H. *From Here to Free Trade*. Chicago: The University of Chicago Press, 1998.

Redding, Stephen, and Venables, Anthony J. "Economic Geography and International Inequality." Centre for Economic Policy Research Discussion Paper No. 2568, 2000.

Rivers, Douglas, and Vuong, Quang H. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39 (1988): 347-366.

Road Atlas of Europe. Chicago: Rand McNally and Co., 1988.

Rose, Andrew. "Do We Really Know That the WTO Increases Trade?" *American Economic Review* 94, no. 1 (March 2004): 98-114.

Rose, Andrew, and van Wincoop, Eric. "National Money as a Barrier to International Trade: The Real Case for Currency Union." *American Economic Review Papers and Proceedings* 91, no. 2 (May 2001): 386-390.

Sapir, Andre. "Trade Benefits under the EEC Generalized System of Preferences." *European Economic Review* 15 (1981): 339-355.

Sapir, Andre, and Winter, Chantal. "Services Trade." In *Surveys in International Trade*, edited by David Greenaway and L. Alan Winters. Oxford, UK: Blackwell, 1994.

Tinbergen, Jan. *Shaping the World Economy*. New York: The Twentieth Century Fund, 1962.

Tomz, Michael, Judith Goldstein, and Douglas Rivers. "Membership Has Its Privileges: The Impact of GATT on International Trade," Stanford University, Dept. of Political Science Working Paper, June 2004.

Trefler, Daniel. "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy." *Journal of Political Economy* 101, no. 1 (February 1993): 138-160.

U.S. Department of the Navy Oceanographic Office. *Distance Between Ports*. H.O. Publication No. 15, U.S. Government Printing Office, 1965.

Vella, Francis, and Verbeek, Marno. "Estimating and Interpreting Models with Endogenous Treatment Effects." *Journal of Business and Economic Statistics* 17, no. 4 (October 1999): 473-478.

Wooldridge, Jeffrey M. *Introductory Econometrics.* Cincinnati, OH: South-Western College Publishing, 2000.

Wooldridge, Jeffrey M. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

World Bank. *World Development Indicators*, CD-ROM. Washington, DC: World Bank, 2001.

World Trade Organization. http://www.wto.org .